Ron Basich opened the session by reviewing various research designs for the study of writing assessment. He drew upon his professional experience as research analyst in the California Chancellor's Office where he has been instrumental in developing a multi-test data base for analyzing the entry level placement and admissions tests for the California State University System.

He delineated three major factors that aid research design: 1) the identification, understanding, and control of variables; 2) the use of accurate measuring devices and the understanding of variance within these instruments; and 3) project management. In order to identify and account for variance factors in research design, one must identify data sources and the impact of these sources on the design. Questions to ask while setting up a design are: 1) Will the data and its analysis prove generalizable? and 2) Is consideration of inter-rater effects built into the design? With the identification and development of appropriate measuring devices, it is essential to know the subject area being assessed. This in turn aids construct validity or the extent to which the measuring device captures the heart of the questions being studied. Further, context validity is strengthened because the many facets of the subject matter studied are identified and their effects accounted for.

While redundancy in test items should be avoided, reliance on one item to measure a particular content question may be insufficient. Related items, therefore, should elicit information that serves to elaborate content area from many different angles. In this way, unambiguous responses are facilitated while duplication across items is avoided. Test instruments should be pilot tested early on in a project in order to ensure adequate time to facilitate improvement of responses to test items. Early checks on items help lessen the probability of external and extraneous variance in responses. The means for testing should be set up and clearly established ahead of time. Some functional issues to consider are the compactness of the test, time allowances scheduled for test administration, means of distributing tests to subjects, and proper coding for essay assignment during analysis.

The key to effective project management is to encourage cooperation among project members. This becomes essential in large-scale research designs in which efficient data collection becomes contingent upon the cooperation of each project member. A further concern of the project manager is to generate the comprehensive cooperation of project participants. Allowing flexibility to enter the design can result in useful feedback from project members and participants. Paying strict attention to the realities of funding and the distribution of funds over time and coordinating funds between grant cycles and federal fiscal years and all contribute to managing a project smoothly.
RESEARCH ON WRITING ASSESSMENT

Speakers: Arthur Applebee, Stanford University
Gordon Brossell, Florida State University

Introducer/Recorder: Dick Worthen, Diablo Valley College

Arthur Applebee began by noting that the Education Department at Stanford University has been studying writing across the curriculum in a number of secondary schools — not just writing in the English Departments. Some tentative observations from this study in progress that Arthur Applebee shared with the group are:

1. Less than half of student writing is done in English classes;
2. About three-fourths of student writing deals with specific ideas presented in class, not with personal experiences;
3. Writing is more often assessed than taught;
4. In most student writing, there is less emphasis on originality and organization and more on accuracy of reporting and on information portrayed;
5. Students seem best at, and favor, writing narratives. They seek ways to work narrative into an assignment that calls for analysis. If it is insisted that the writing be analytical, their achieved fluency usually falters.

Applebee noted that there are many differences within a faculty as to what good writing is — differences as to why and how we teach writing. This makes it difficult to develop school-wide standards for assessment and exposing a major problem of controlling how pedagogical method, student perception, and writing assessment interact.

Gordon Brossell pointed out that testing often drives curriculum. Because the statewide Florida writing test, The College-Level Academic Skills Test in Writing, calls for a short expository essay, a good many community-college programs have begun to emphasize the writing of fifty-minute compositions (i.e., practicing for the test). If writing assessments like this one were geared to writing processes rather than to short salvos of writing production, they might very well stimulate the writing teacher (and writing programs) to pay more attention to the craft of composing — to generating ideas, gathering data, drafting, revising, editing, and proofreading. Assessments modeled after the composing process would get the message across that this is what happens in good writing classes and that this should happen in all classes.

RESEARCH DESIGNS (continued)

and 1500 essays were collected for analysis. Each essay was holistically scored and inter-rater reliability checks were conducted throughout the scoring procedure. Two discourse features were analyzed. One category, development and focus, consisted of the ability to state generalities and support those generalities. Another, high level mechanics, included evidence of proper sentence embedding and varieties of sentence constructions. This study is still in progress, so no final report was made. However, it proved extremely useful as a model of how to set up an effective research design and helped illustrate issues to consider in managing a complex research project.

Brossell reminded us that the means of testing can and should affect instruction. A large question, then, is how in assessing writing do we build into writing prompts the stimuli that evoke what we desire from students? And how do we get those teaching writing to coordinate curriculum and programs in such a manner as to support and prepare the student being tested and make rational the assessing process? We must move carefully to assure that the cause-effect relationship is asserted in the right direction. Brossell suggested that we should start with the questions "Why do we teach writing and what skills in writing comprise the qualities in the writer we want the schools to produce? How can we achieve such consensus across the curriculum?"

Brossell has surveyed the area of student response to the writing prompt. He notes three variables in writing assessment: topic variables, writer variables, and procedural variables. How do we muster our efforts as teachers and testers to achieve feedback that serves the purposes we have created our curriculum for? Brossell gave us several generalizations that deserve our attention but may leave a slightly uneasy as we seek some larger verities for pinning down an admittedly elusive field:

1. Small differences in wording within the same general framework seem to make little difference in student responses.
2. Topics with low cognitive demands and high experiential demands elicit higher scores from readers.
3. The problem of unequal familiarity with the topic can usually be overcome by supplying ample information.
4. A writing prompt calling for an argument rather than a narrative is more difficult to respond to, especially among young respondents.
5. Prompts that are at least moderately specified (rather than open-ended) elicit more focused and better organized essays. This is more important in a timed essay.

Brossell concluded by noting that Alan Purves and his colleagues at the Curriculum Lab of the University of Illinois have been working on a model of the composition assignment which sets forth fifteen dimensions of a writing assignment: instruction, stimulus, cognitive demand, purpose, role, audience, content, discourse, specification, tone and style, preparation, length, format, time, number of drafts, and criteria for evaluation. The categories are intended to give test-makers and teachers a set of tools for "adjusting" writing topics.

PROFICIENCY TESTING (continued)

a course or an exam, similar to that given to satisfy the Subject A requirement, but more sophisticated. They too are evaluated by the writing program staff, which reads them holistically. At Davis, a score of 8 (out of a possible 12) is necessary to pass, and the percent passing ranges from the high 40's to the mid 60's. At Irvine, where 11 or better is a passing score, the pass rate is 8-10%. In both universities, this test functions as a challenge exam that exempts students from the upper-division requirement; those who fail the exam must take an upper-division composition course.

Soltwisch, Lothamer, and Gadda concluded by urging faculty who participate in test development to be concerned with appropriate, fair, and legally-defensible assessments of writing skill.